

# Trans-spliced Cas9 allows cleavage of *HBB* and *CCR5* genes in human cells using compact expression cassettes

Eli J. Fine, Caleb M. Appleton, Douglas E. White, Matthew T. Brown, Harshavardhan Deshmukh, Melissa L. Kemp, and Gang Bao

## Supplementary Information

### 1. Supplementary Methods

Supplementary Method M1. PCR primers for T7E1 Assays ..... 2

### 2. Supplementary Figures

Supplementary Figure S1. Potential layout of future multi-plexed trans-splicing system ..... 3

Supplementary Figure S2. Initial test of DnaB and GyrA intein systems..... 3

Supplementary Figure S3. Representative T7E1 gel of nuclease activity..... 4

### 3. Supplementary Data

Supplementary Data D1. Amino acid sequences of trans-splicing components ..... 5

Supplementary Data D2. Plasmid sequences of trans-splicing components ..... 6

### 4. Supplementary References

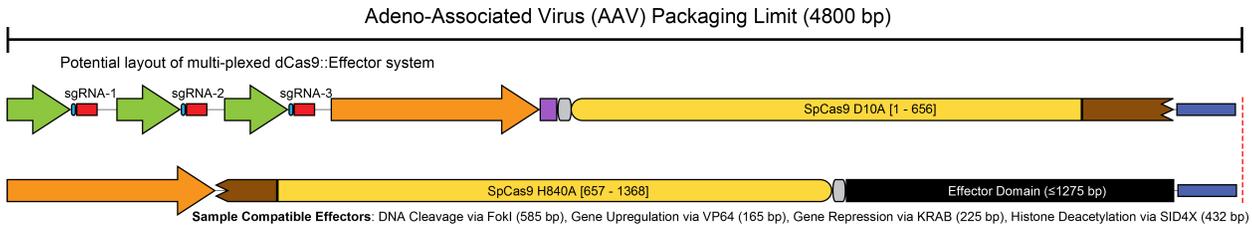
## 1. Supplementary Methods

### Supplementary Method M1. PCR primers for T7E1 Assays

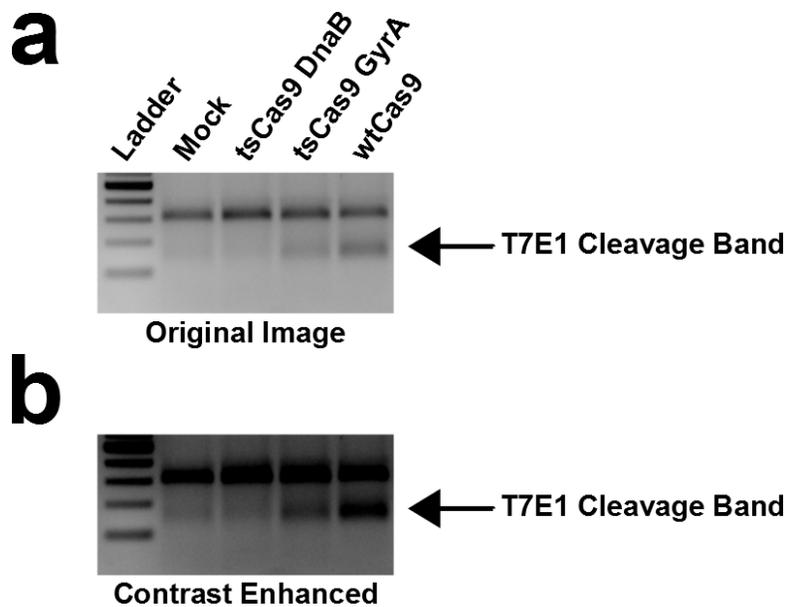
Gene	Forward Primer	Reverse Primer
<i>HBB</i>	CTGGAGACGCAGGAAGAGATCC	GCAATCATTCGTCTGTTTCCCATTC
<i>CCR5</i>	AGTCGACACTGCACAGGGTGAACAAGATGG	GCATAGATCGACCACCCCAAAGGTGACCGT

All DNA sequences are written 5'→3'. All oligos were ordered from Eurofins MWG Operon. The CCR5 primers have barcode sequences on the 5' ends because they were originally designed for multiplexed DNA sequencing; this has no effect on the T7E1 assay.

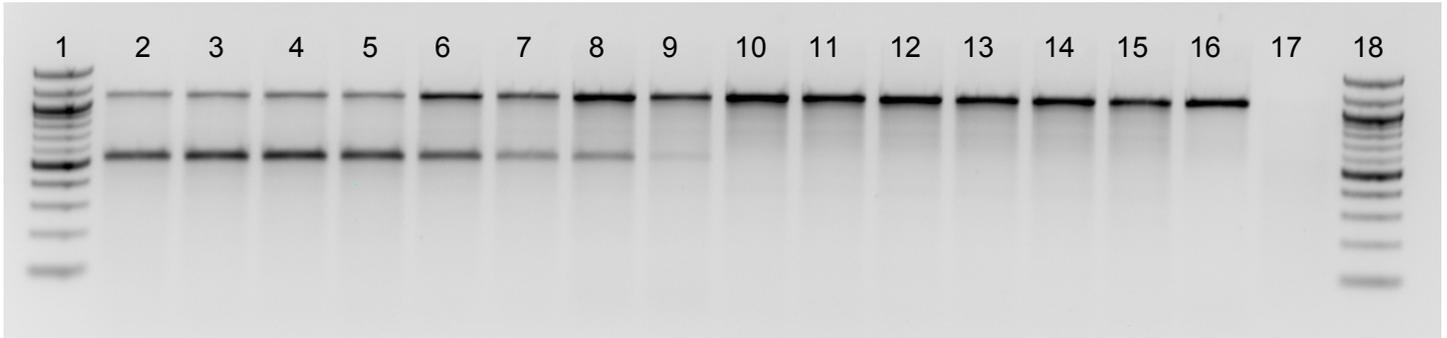
## 2. Supplementary Figures



**Supplementary Figure S1. Potential layout of future multi-plexed trans-splicing system** using synergistic sgRNAs and an effector fused to dCas9 to allow gene regulation or epigenetic modification.



**Supplementary Figure S2. Initial test of DnaB and GyrA intein systems.** To initially determine the feasibility of using the different intein systems, DnaB and GyrA trans-splicing Cas9 systems were transfected in triplicate into HEK-293T cells along with wild-type Cas9. All plasmids contained the guide strand targeted to *HBB* (See **Figure 3a, Supplementary Method M1**). Shown here is a representative gel (a) from the T7E1 assay and a version of the image with enhanced contrast to highlight the cleavage band (b). Because the DnaB system did not induce gene modification at a rate appreciably above the background signal in mock-treated cells, it was not considered further. We recognize the unwanted presence of a faint band appearing in the mock treated cells; this initial test was performed with an earlier set of PCR primers and a new set of primers (used in all other experiments) resolved this issue.



**Supplementary Figure S3. Representative T7E1 gel of nuclease activity.** T7E1 reactions were loaded into 2% agarose gels cast with ethidium bromide. This gel analyzed one replicate of the dosing experiment of the R-3 guide strand targeting *HBB*. The PCR amplicon is ~1100 bp and the two T7E1 cleavage products both co-localize at ~550 bp. Activity was observed for cells transfected with the wtSpCas9 plasmid and with both components of tsSpCas9, but no activity was observed (as expected) for control reactions containing only pUC, only one of the tsSpCas9 components, or Cas9 systems guided to *CCR5* and not to *HBB*.

**Gel Lanes (from left to right):**

- |                            |                                                    |
|----------------------------|----------------------------------------------------|
| 1 – 100 bp ladder          | 10 – N-terminal tsSpCas9 R-3 only                  |
| 2 – wtSpCas9 R-3 Full Dose | 11 – C-terminal tsSpCas9 R-3 only                  |
| 3 – wtSpCas9 R-3 ½ Dose    | 12 – N-terminal tsSpCas9 R-30 ( <i>CCR5</i> ) only |
| 4 – wtSpCas9 R-3 ¼ Dose    | 13 – C-terminal tsSpCas9 R-30 ( <i>CCR5</i> ) only |
| 5 – wtSpCas9 R-3 1/8 Dose  | 14 – wtSpCas9 R-30 ( <i>CCR5</i> ) Full Dose       |
| 6 – tsSpCas9 R-3 Full Dose | 15 – tsSpCas9 R-30 ( <i>CCR5</i> ) Full Dose       |
| 7 – tsSpCas9 R-3 ½ Dose    | 16 – pUC only                                      |
| 8 – tsSpCas9 R-3 ¼ Dose    | 17 – No template PCR control                       |
| 9 – tsSpCas9 R-3 1/8 Dose  | 18 – 100 bp ladder                                 |

### 3. Supplementary Data

#### Supplementary Data D1. Amino acid sequences of trans-splicing components

Intein domains underlined.

##### >tsCas9-C

MGKPEFAPTTYTVGVPLVRFLEAHHRDPDAQAIADDELTDGRFYAKVASVTDAGVQPVYSLRVDTADHAFITNGFVSHNTGWGRLSRKLINGIRDKQSGKTIIDFLKSDGF  
ANRNFQMQLIHDDSLTFKEDIQKAVSGQDLSLHEHIANLAGSPAIAKKIGLQTVKVVDELVKVMMGRHKPENIVIEMARENQTQKGQKNSRERMKRIIEGKELGSQLIKEHP  
VENTQLQNEKLKLYLQNGRDMYVQELDINRLSDYDVHIVPQSFLKDDSDI DNKVLTRSDKNRGRKSDNVPSEEVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGSELD  
KAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVIITLKSCLVSDFRKDFQFYKVIINNYHHAHDAYLNAVVGITALIKKPKLESEFVYGDYKVDVRKMIK  
SEQEI GKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLVSMQVNI VVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDS  
PTVAVSVLVAVKVEKSKLKS SVKELGITIMERSSEFEKNPIDFLEAKGYKEVKKDLI IKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFYLLASHYEKLKGS  
PEDNEQQLFVEQHKLHLYDEIIEQISEFSKRVI LADANLDKVL SAYNKHHRDKPIREQAENI IHLFTLTLNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQISITGLYETRI  
DLSQLGGDKRPAATKKAGQAKKKK

##### >tsCas9-N

MDYKDHGDYKDHDI DYKDDDDKMAPKKRKYVIGVPAADKKYSIGLDIGTNSVGVAVITDEYKVPSSKFKVLGNTDRHS IKKNLIGALLFDSGETAEATRLKRTARRRYT  
RRNRICYLQEIFSNEMAKVDDSFHRLLESFLVEEDKKHERHPI FGNIVDEVAYHEKYPTIYHLRKKLVDS TDKADLRILIYALAHMIKFRGHFLIEGDLNPDNSDVDKLF  
IQLVQTYNQLFEENPINASGVDAKAIL SARLSKSRLENLIAQLPGEKKNLFGNLI ALSGLTPNFKSNFDLAEADAKLQLSKDYDDDLNLLAQIGDQYADFLAANKNS  
DAILLSDILRVNTEITKAPLSAMIKRYDEHHQDILLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQE EYFKIPKILEKMDGTEELLVKLNREDLLRKQRTFDNGS  
IPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRI PYVYGPLARGNSRFAMWTRKSEETIPWNFEVVVDKGAQAQSFIERMTNFDKNLPNEKVL PKHSLLYEYFTVY  
NELTKVKYVTEGMRKPAFLS GEQKKAIVDLLFKTRNKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLKI IKDKDFLDNEENEDILEDIVLTLTLFEDREMIE  
ERLKYAHLFDDKVMKQLKRRRYCITGDALVALPEGESVRIADIVPGARPNSDNAIDLKVLDRHGNPVLADRLFHSGEHPVYTVRTEVGLRVGTANHP LLLCLVDVAGVPTL  
LWKLIDEIKPGDYAVIQRSAFSVDCAFAR

##### >tsCas9-Nick

MYPYDVPDYASPKKKRVEASDKKYSIGLAIGTNSVGVAVITDEYKVPSSKFKVLGNTDRHS IKKNLIGALLFDSGETAEATRLKRTARRRYTRRNRICYLQEIFSNEMAK  
VDDSFHRLLESFLVEEDKKHERHPI FGNIVDEVAYHEKYPTIYHLRKKLVDS TDKADLRILIYALAHMIKFRGHFLIEGDLNPDNSDVDKLF IQLVQTYNQLFEENPINAS  
GVDAKAIL SARLSKSRLENLIAQLPGEKKNLFGNLI ALSGLTPNFKSNFDLAEADAKLQLSKDYDDDLNLLAQIGDQYADFLAANKNSDAILLSDILRVNTEITKAP  
LSAMIKRYDEHHQDILLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQE EYFKIPKILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQED  
FYFPLKDNREKIEKILTFRI PYVYGPLARGNSRFAMWTRKSEETIPWNFEVVVDKGAQAQSFIERMTNFDKNLPNEKVL PKHSLLYEYFTVYNELTKVKYVTEGMRKPAFL  
SGEQKKAIVDLLFKTRNKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLKI IKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKYAHLFDDKVMKQLK  
RRRYCITGDALVALPEGESVRIADIVPGARPNSDNAIDLKVLDRHGNPVLADRLFHSGEHPVYTVRTEVGLRVGTANHP LLLCLVDVAGVPTLLWKLIDEIKPGDYAVIQRS  
AFSVDCAFAR

#### Supplementary Data D2. Plasmid sequences of trans-splicing components

##### >tsCas9-C

gagggcctatttccccatgatttccttcataatgtgcataacgatacaagcgtgttagagagataaattggaattaatttgactgtaaacacaaagatatttagtacaataatcgt  
gacgtagaaagtaataat tcttgggtagtttgcagttttaaataatgttttaaataatggactatcatatgctaccgtaactgaaagtatttcgatttcttggccttata  
tatcttGTGGAAAGGACGAAACACCggGTCTTCAagctGAAGACctgttttagagctGAAAtagcaagttaaaataagcgttagtccttattcaactgaaaagtggcac  
cgagtcggtgctTTTTggttagagctagaaatagcaagttaaaataagcgttagtccttattcaactgaaaagtggcac  
cataacttacggtaaatggcccctggctgaccccaacgaccccccccattgacgtcaatagtaaacgcaatagggactttccattgacgtcaatgggtggagatt  
acggtaaacctggcactggcagtagatcaagtgatcatatgcaagtagcaccctctattgacgtcaatgacgtaaatggcccctggcattGtgcccagtagcatgacc  
ttatgggactttcctacttggcagtagatctacgtatttagtcatcgctattaccatggctgaggtgagccccagcttctgctcactctccccctccccac  
cccaattttgtaattttatttttttaattttttgtgcagcagtagggggggggggggggggggggggcgcgccagggcggggcgggggcgaggggggggggc  
gagggcgagaggtggcgggcgagcacaatcagagcgggcgctccgaaaagtttcttattgagcagggcgggcgggcgcccataaaaaagcgaagcgcgggcgggcg  
ggagtcgtagcagctgcccctggcagtagcaccctgcccctggcagcagctcgcgcggcctggcggcggcctctgactgacccgcttactcccacaggtgagcgggcggc  
ccttctctcgggctgtaattagctgagcaagaggttaaggttttaagggatggttgggtggtggggtatataatgtttaataactggagcactgctgaaatcactttt  
ttcaggttGGaccggtgcccaccATGGGAAAACCGAAATTTGCCCAACTACATATACGTGCGCGTCCAGGGCTGGTTCGATTCCTGGAAGCCACCACAGGGATCCAGAC  
GCTCAGGCAATCGCAGACGAGCTGACCGAGGGAGGTTTTATTATGTCAAGCTGCCTGTGTGAGCCGATCGAGGCGTACAGGCGTACAGGCGTGTGTATCTCTAGGCTGGACAGCCG  
ACCATGCTTTTATTACCAATGGCTTCGTGTAGTCACAATACCTGTGCGGAGGCTGAGCCGGAAGCTGATCAAGGCGATCCGGGACAGAGCTCCGGCAAGCAATCGTGA  
TTCTCTGAAGTCCGACGGCTTCGCCAACAGAACTTCATGCAGCTGATCCACGACGACAGCCTGACCTTAAAGAGGACATCCAGAAAGCCAGGTTCCGGCCAGGGCGAT  
AGCTGCACGAGCACATTCGCAATTCGCGCGGAGCCCGCCATTAAAGAGGCGATCCTGCAGACAGTGAAGGTGGTGACGACGAGCTCGTGAAAGTGTGGCCGCGCACAAAGC  
CCGAGAATCGTGATCGAAATGGCCAGAGAGAACCAGACCACCCAGAAGGGCAGAGAAGACAGCCGCGAGAGAATGAAGCGGATCGAAGAGGGCATCAAAGAGCTGGGCAG  
CCAGATCCTGAAAAGAACCCCGTGGAAAACACCCAGCTGCAGAACGAGAGCTGTACTGTACTACCTGCAGAAATGGCGGGATATGTACGTGGACCAGGAACCTGGACATC  
AACCAGCTGTCCGACTACGATGTGGACCATATCGTGCCTCAGAGCTTTCTGAAGGACGACTCCATCGACAACAGGTGCTGACCAGAAGCAGACAAGAACCCGGGCAAGAGCG  
ACAACGTCGCCCTCGAAGAGGTGTAAGAAGTGAAGAACTAGCGCGGACGTGCTGAAAGCACAAGCTGATTACCAGAGAAGTTCGACAACTGACCAAGGCCGAGAG  
AGGCGGCTGAGCGAAGCTGGATAAGGCCGGCTTCATCAAGAGACAGCTGGTGGAAACCCGGCAGATCAAAAGCACGTTGGCACAGATCCTGGACTCCCGGATGAACACTAAG  
TAGCGAGAGAATGACAACTGATCGGGAAGTGAAGTGAATCACCTGAGTCAAGCTGGTTCGATTCGGAAGAGGATTTCCAGTTTCCAGAAAGTGGCGGAGATCAACA  
ACTACCACCACGCCACGACGCCCTACCTGAACGCCCTGCTGGGAACCGCCCTGATCAAAAAGTACCTTAAGCTGGAAAGCGGATTCGTGTACGGCAGTACAAAGGTGTACGA  
CGTGGGGAAGATGATCGCAAGAGCGAGCAGGAAATCGGCAAGGCTACCGCAAGTACTTCTTCTACAGCAACATCATGAACCTTTTTCAGAGCCGAGATTACCTGGCCAAC  
GGCGAGATCCGGAAAGCGGCTTCATCGAGACAACCGGCAACCCGGGAGATTCGTGTGGGATAAGGCCGGGATTTTGGCACCGTGGGAAAGTGTGAGCATGCCCAAG  
TGAATATCGTGAAGAAGCAGGTCAGGTCAGAGCCGCTACTCTGCCCCAAGAGAACCGATTAAGCTGATCGCGGAGAGCAGTATGATCGCGGAGAGAGGACTGGGACATTA  
GAGATACGGCGCTTCGACAGCCCCACCGTGGCCTATCTGTGCTGGTGGTGGCCAAAGTGGAAAAGGGCAAGTCCAAGAACTGAAGAGTGTGAAGAGCTGCTGGGGATC  
ACCATCATGGAAGAAGCAGCTTCGAGAAGAACTCCATCGACTTCTGGAAGCAAGGGCTACAAAGAAGTGAAGAAGGACCTGATCATCAAGCTGCCTAAGTACTCCCTGT  
TCGAGCTGGAAAGCGCCGGAAGAGAATGCTGGCCTTCGCCGCGAAGCTGCAAGAGGAAACGAACTGGCCCTGCCCTCCAAGATATGTGAACCTTCCTGTAGCTGGCCACCA  
CTATGAGAAGTGAAGGCTCCCGGAGTAATGAGCAGAAACAGCTGTGTTGGAACAGCACAAGCACTACCTGGAGAGCAGTATCATCGAGCAGATCAGCAGATCTCCCAAG  
AGAGTATCCTGGCCGAGCTAATCTGGACAAAGTGTGTCGCCCTACAACAAGCAGCGGATAAGCCATCAGAGAGCAGGCGGAGATAATCATCCAACCTGTTTACCTGA  
CCAATCTGGGAGCCCTGCCGCTTCAAGTACTTTGACACCACCATCGACCGAAGAGGTACACCAGCACAAAGAGGTGCTGGAGCGCCACCTGATCCACCAGAGCATCAC  
CGCCCTGTACGAGACCGGATCGACTGCTCAAGTGGGAGGCGCAAAAGCCGGCGCCAGGCAAAAAGGCCGCGCCAGGCAAAAAGGAAAGGtaagaat tctTAGAGCTC  
GCTGATCAGCTCAGCTGCTGCTTCTAGTTGCCAGCATCTGTGTTGCCCTCCCGTGCCTTCTTACCTGGAAGGTGCCACTCCCCTGTCTTCTTAATAAAA



cttcccggcaacaattaaactagactggatggaggcggataaaagtgcaggaccacttctgcgctcggccctccggctggctgggtttatgtgtataaaactgagaccgggtga  
gcgtggaagccgggtatcattgacagactggggccagatggttaagccctccgtagtctatctacacagcggggagtcaggcaactatggatgaacgaataagacag  
atogctgagatggtgctcactgactaagcaatggtaactgactgcagacaagttactcatatactttagatgtatataaaactctatctttaaagatctag  
tgaagatccttttgataatctcatgacccaaaatcccttaacgtgagtttctgctccatgagctagcctagaccocctgagaaaagatcaaggatctctcttgagatccttttt  
tctgcgctaatctgctgcttgcacaacaaaaaaccccgctaccagocgggtgtttgtttgocggatcaagagctaccaactcttttccgaaggttaactggctcagcaga  
gycagataccaataactgtccttctagtgtagccgtagtttagggccaccactcaagaactctgtagcaccgctacatacctcgctctgtaactcctgttaccagtggtg  
ctgccagtggcgataagctgcttcccggttggactcaagagctagttaccggataagcggcagcggctgggctgaacggggggtctgctgcacacagccagcttgg  
gcaacagcactacccggaactgagatacctacagcgtgagctatgagaagcgccacgcttccggaaggagaaaaggcggagaggtatccggtaagcggcagggctggaa  
ggagagcgcacaggggagctccagggggaacgcctggtatctttagctctgctgggttccgccacctctgactgagcgtcgatcttctgtagctcgtcagggggg  
ggagcctatgaaaaaacgcccagcaacgccccttttacggttccctggccttttctgctggccttttctgctcaatgTCCTGCAGGCGAGCTGCGCGCTCGCTCGCTCACTGAGG  
CGCCGGGCAAGCCCGGGGCTGGGGACACTTTGGTGCCTGGCCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAACCTCCATCACTAGGGGTTCCTAcatgt

**>tsCas9-Nick**

gagggcctatttcccattgatccttctcatatttgcataacgatacaagggctgtagagagataattggaattaatttgactgtaaacacaaaagatatttagtcaaaaatcgt  
gacgtgaaagtaataatttctgggtagtttgcagtttaaaatattgttttaaaaggactatcatatgcttaccgtaacttgaagatttctgatttcttggctttata  
tatcttGTGGAAGACGAAACACCggtCTTTCgaGAGACCTgttttagctgataGAAAtagcaagtttaaaataaggttagtcggttcaacttgaaaaagtggaaccgag  
tcggtgctTTTTTgttttagagctagaaaatagcaagttaaaataaggttagtcggtTTTTAgcggctgcccacttctgcagacaaaaggctctagaggtaccctgtacata  
acttacggtaaatggcccgctggctgaccgcccacgaccccggccattgacgtcaatagtaaacgccaatagggactttccattgacgtcaatgggtggagtattttacg  
taaactgcccacttggcagatcaatgtagtgcacagtaagcagccttattgacgtcaatgacgggtaaaatggcccctggctgattGtggccagtagtgcactt  
gggacttctcacttggcagtagatcactgtagttagtcatcgtattaccatggctgagggtagccccagcttctgcttcaacttccccatcccccccccccc  
aattttgtattttatatttttaattttttgtgtagcagatggggggggggggggggggggcgcgcgcccagggcggggcggggcggggcggggcgaggggcggggcggggcggggcggg  
cggagaggtgcccggcgaagcaatcagagcggcgcgctccgaagtttcttttagcggagcggcggcgggcgccctataaaaagcgaagcgcggcgggggggag  
tcgctgcgagcgtgcttccgcccgtgcccgcctcgcggcgcctcgcggcggcggcggcggctctgactgaccgcttactcccacaggtgagcgggggggacggcctt  
ctctcggcgtgataatgctgagcaaggttaaggtttaaaggttagttgggttgggtatataagtttaattacctggagcagctgtaactggaatccttttttca  
ggttGGaccggtgccaccATGTACCATACGATGTTCCAGATTACGCTTCGCCGAAAGAAAAGCGCAAGGTCGAAGCGTCCGACAAGAAGTACAGCATCGCCCTGGCCATCG  
GCACCAACTCTGTGGGGTGGGGCGTGTACCCGACGAGTACAAGGTGCCAGCAAGAAATCAAGGTGCTGGCAACACCAGCCGGCAGACGATCAAGAAGAACCTGTATCGG  
AGCCCTGCTGTTCGACAGCGGGAACAGCCGAGCCAGCCGCTGAAGAGAACCCAGAGAGATACACCGAGGAAAGAACCCGATCTGCTATCTCGAAGAGATCTTC  
AGCAACAGAGATGGCCAAAGTGGACAGCTTCTCCACAGCTGGAAGACTTCTCCTGAGTAAGAGTAGAAGAAGCAGCAGGACCCCATCTTCGCGAACACTCGTGG  
ACGAGGTGGCCTACCACGAGAAGTACCCACCATCTACCACCTGAGAAAGAACTGGTGGACAGCACCCGACAAGGCCAGCTGCGGCTGATCTATCTGGCCCTGGCCACAT  
GATCAAGTTCGCGGGCACTTCTGATCGAGGCGGACCTGAACCCCGACAACAGCGAGCTGGACAAGCTGTTTCATCCAGCTGGTGCAGACCTACAACCAGCTGTCGAGGAA  
AACCCCATCAACGCCAGCGGGCTGGAGCCAAAGGCCATCTGTCTGCCAGACTGAGCAAGAGCAGCAGCGGTGGAAATCTGATCGCCAGCTGCCCGGCGAGAAGAAGT  
GCCTGTTTCGGCAACCTGATTCGCTGAGCCCTGGGCTGACCCCAACTCAAGAGCACTTGCACCTGGCCGAGGATGCCAAACTGCAGCTGAGCAAGGACACTACGACGA  
CGACCTGGACAACCTGCTGGCCAGATCGGCGACCGTAGCCGACCTGTTTCTGGCCGCAAGAACCTGTCCGACGCCATCTGCTGAGCGACATCTGAGAGTGAACACC  
GAGATACCAAGCCCGCTGAGCGCTCTATGATCAAGAGATACGAGCAGCACCCAGGACTGACCTGCTGAAAGCTCTCGTGGCGCAGCAGCTGCCTGAGAAGTACA  
AAGAGATTTCTTCGACCAGAGCAAGACCGGTACGCCGGCTACATTGACGGCGGAGCCAGCAGGAAAGAGTCTACAAGTTCATCAAGCCCATCTTGAAAGAGATGGACGG  
CACCGAGGAACTGCTCGTGAAGCTGAACAGAGGACCTGCTCGGAAGCAGCGGACCTCGCAAGCAGCGCAGCATCCCCACAGATCCCTGGGATACCGCTCGCACCTAG  
CTGCGCGGCGAGGAAATTTTACCATTCTGAAAGGACAACCGGAAAGATCGAGAAGATCCTGACCTTCCGATCCCCTACTACTGTTGGCCCTCTGGCCAGGGGAAACA  
GCAGATTGCTGGATGACCCAGAAAGAGCGAGGAAACCATCACCCCTGGAATCTCGAGGAAAGTGGTGGACAAGGGCGCTTCCGCCAGAGCTTCATCGAGCGGATGCCAA  
CTTCGATAAAGAACTGCCAACCGAGAAGGTGCTGCCAAAGCAGCCGCTGTGTACGAGTACTTACCCGTGTATAACAGAGCTGACCAAGTGAATACTGTACCCAGGGAAATG  
AGAAAGCCCGCTTCTTCCAGCAGGAGCAAGAAAGGCCATCTGGAAGTGTGACTGCTTCAAGCAACCCGAAAGTGAACCTGAAAGAGTGAACCTGGGAGGACTCTCAAGAAA  
TCGAGTGTTCGACTCCGTGGAATCTCCGCGTGGAAAGTACGTTCAACGCCCTCCCTGGGCACATACCAGATCTGCTGAAAATTAACAAGGACAAGGACTTCTGGACAA  
TGAGGAAAACGAGGACATCTGGAAGATATCGTGTGACCCCTGACACTGTTTGGAGGACAGAGAGATGATCGAGGAACGGCTGAAAACCTATGCCCACCTGTTTCGACGACAAA  
GTGATGAGCAGCTGAAGCGGCGGAGATACTGTATCACCGTGTGCTGGTGGCGCTGCCCGAAGGCAATCAGTGCATCGCTGACATCGTCCCGGGGCGCCCGCTTA  
ATAGCGATAACGCCATCGACTGAGGCTGAAGGTTGCTGACCGCCACGGCAATCCTGTCTGGCCGATGACTCTCCAGTGGTGAACCTCCCTGATACCGCTCGCACCTAG  
AGGCCCTGAGAGTGAACAGGCACAGCCAACCCCGCTTTTGTGTCTGGTGGATGTTGCCGGCTGCCACCCTGCTGTGGAACCTGATTGATGAGATAAAAACCGGAGACTAT  
GCAGTGTCCAAAGTCCGCTTGTAGTGTGGATGCGCTGGCTTCGCTCGCTcaagaattcctTAGAGCTCGCTGATCAGCCCTGACTGTGCCTTCTAGTTGCCAGCCATCTGT  
TGTTTGGCCCTCCCCGCTTCTTGAACCTGGAAGGTGCCACTCCACTGCTCTTCTTAATAAAAATGAGGAAATGCACTGCAATGTCTGAGTAGGTTGATCTTAT  
CTGGGGTGTGGGTGGGCGAGCAGCAAGGGGAGGATTGGGAAGAGAAATAGCAGGCTGCTGGGGAGcggcggcaggaaaccctagtgatggagttggccactccctctc  
tgcgcgctcgctcgctcactgagccggcgagcaaaaaggctgcccgaacgcccggccttggccggggcctcagtgagcagcggagcggcgagctgcttggcagggggcct  
gatggggtatcttctccttacgcatctgtgcggtatctcacacogcatcgtcaaagcaacctatgtagcgcocctgtagcggcgcataaagcggcggggtgtgggtgta  
cgcgagcgtgacccgctacacttgcacgcgcccctagcggccgctccttcccttctcctccttctcggccagcttgcggcttccccgctcaagcttaaatcgggg  
gctccctttagggttccgatttagcttttaggcacctcgacccccaaaaactgatttggtgtagtgctacgtaggggccatgoccttagacagcgtttttggcct  
ttagcgttggagtcacagttctttaaagtgactcttctgtaaacactggaaacacactcaaccctattctcgggctattcttggattttaaaggttttgccgatttctg  
cctattggttaaaaaatgagctgatttaacaaaatttaacgogaatttaacaaaataaaacgtttacaatttttaggtgactcctcagtaaatctgctctgatggc  
atagttaaagcaagcccgacaccggcaaacaccgctgacgcgcccgtgacgggttctgctgcggcctcagctacagacaagctgtgacgcttccgggagctgcag  
tgtcagagttttcaactgctcagccgaacccgagacgaagggcctctgatacgcctatttttaggttaatgtagtataataatggtttcttagactcaggtg  
gctatttggggaaatgtgcccagccccctatttggttattttcaataatcattcaataatgtagtcctcagtagagacaataagctgataaatgtaaatgta  
aaaaggaagagtagtagtattcaacattccggtgctgcccctattccctttttgocgcatttgocctcctggttttctgctcacccagaaacgctggtgaaagtaaagat  
gctgaagatcagttgggtgcacgagtggtttacatcgaactggtatcaacacgctgaagatccttagagatttgcggccgaaagcgtttccaatgatgagcacttta  
aagttctgctatgtggcggtattatccgatttagcgcgggcaagagcaactcgtgcgcgatacactattctcagaatgactggttgagtagtaccagctcacaga  
aaagctcttacggatgtagcagtagaagaaattagcagtgctgcatgccataaccatagtagtacaactcggcccaacttctctgacaacagatcggaggaccgaagag  
ctaaccgcttttttgacaacatggggatcagtagtaactcgccttgatctgggaacccggagctgaatgaagccataccaaacgacgagcgtgacaccacagatgctgtag  
caatggcaacaagcttgcgcaaatcttaactggcaactactcttagcttcccggcaacaataatagactgtagggggcgaataaagttgcaaggaccactctg  
ctggcccttccggtggctggtttatgtgataaaatctggagcctgtagcgtggaagccgctgtagttagcagctgggccagatggtgtagccctccgtatcgt  
gtatctacacgacgggtaggcaactatggatgaacgaataagacagactcgtgagataggtgctcactgattaagcattgttaactgttaagcagcaaggtttactcat  
atatactttagattgatttaaaactcatttttaattaaaagagcttaggtgaagatcctttttagataatctcatgacaaaatcccttaacgtgagtttctgctccactg  
agcgtcagaccggtagaaaagatcaaaaggatcttctgagatcctttttctgocgtaattctgctgctgcaaaaacaaaaccaccgctacacggggtgtttgtttg  
ccgtagcaagagtagtccaactcttttccgaagtaactggtctcagcagagcgcagataccaataactgctctttagttagcgtagtttagggccaccactcaagaact  
ctgtagaccgctacatacctcgtctgtaactcctgttaaccagtgctgcccagtgagcaataagctgctctaccgggttggactcaagacagatagttcaagataa  
ggcgcagcggctgggctgaacggggggttctgtagcacacagcccagctggagcgaacgacctacccgaactgagatacctacagcgtgagctatgagaagcggccacgct  
cccgaaggggaaagcggagcaggtatccggttaagcggcagggctcgaacaggaagcgcacgagggagcttccagggggaacgctggatctttatagctctgtcgggt  
ttcgcacactctgactgagcgtcgatcttctgtagctcgtcagggggggggagcctatggaaaaacggcagcagcggccttttaccggttctcgtgcttctgtgct  
tttctgctcaatgTCCTGCAGGCGAGCTGCGCGCTCGCTCACTGAGGCGCCCGGGGCTGGGGCACTTTGGTGCCTGGCCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAACCTCCATCACTAGGGGTTCCTAcatgt

#### 4. Supplementary References

1. Hsu, P.D. et al. DNA targeting specificity of RNA-guided Cas9 nucleases. *Nature Biotechnology* **31**, 827-832 (2013).
2. Cradick, T.J., Fine, E.J., Antico, C.J. & Bao, G. CRISPR/Cas9 systems targeting  $\beta$ -globin and CCR5 genes have substantial off-target activity. *Nucleic Acids Research* **41**, 9584-9592 (2013).
3. Kurpiers, T. & Mootz, H.D. Site-Specific Chemical Modification of Proteins with a Prelabelled Cysteine Tag Using the Artificially Split Mxe GyrA Intein. *ChemBioChem* **9**, 2317-2325 (2008).
4. Martin, D.D., Xu, M.Q. & Evans, T.C., Jr. Characterization of a naturally occurring trans-splicing intein from *Synechocystis* sp. PCC6803. *Biochemistry* **40**, 1393-1402 (2001).
5. Brenzel, S., Kurpiers, T. & Mootz, H.D. Engineering Artificially Split Inteins for Applications in Protein Chemistry: Biochemical Characterization of the Split Ssp DnaB Intein and Comparison to the Split Sce VMA Intein†. *Biochemistry* **45**, 1571-1578 (2006).
6. Nishimasu, H. et al. Crystal Structure of Cas9 in Complex with Guide RNA and Target DNA. *Cell* **156**, 935-949 (2014).
7. Humphrey, W., Dalke, A. & Schulten, K. VMD: visual molecular dynamics. *Journal of molecular graphics* **14**, 33-38, 27-28 (1996).
8. Esvelt, K.M. et al. Orthogonal Cas9 proteins for RNA-guided gene regulation and editing. *Nature Methods* **10**, 1116-1121 (2013).
9. Altschul, S.F. et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic acids research* **25**, 3389-3402 (1997).
10. Tesson, L. et al. Knockout rats generated by embryo microinjection of TALENs. *Nature Biotechnology* **29**, 695-696 (2011).
11. Hockemeyer, D. et al. Genetic engineering of human pluripotent cells using TALE nucleases. *Nature Biotechnology* **29**, 731-734 (2011).
12. Doyon, Y. et al. Enhancing zinc-finger-nuclease activity with improved obligate heterodimeric architectures. *Nature Methods* **8**, 74-79 (2011).
13. Christian, M.L. et al. Targeting G with TAL Effectors: A Comparison of Activities of TALENs Constructed with NN and NK Repeat Variable Di-Residues. *PLoS ONE* **7** (2012).
14. Perez, E.E. et al. Establishment of HIV-1 resistance in CD4+ T cells by genome editing using zinc-finger nucleases. *Nat Biotech* **26**, 808-816 (2008).
15. Pattanayak, V., Ramirez, C.L., Joung, J.K. & Liu, D.R. Revealing off-target cleavage specificities of zinc-finger nucleases by in vitro selection. *Nat Meth* **8**, 765-770 (2011).
16. Gupta, A., Meng, X., Zhu, L.J., Lawson, N.D. & Wolfe, S.A. Zinc finger protein-dependent and -independent contributions to the in vivo off-target activity of zinc finger nucleases. *Nucleic Acids Research* **39**, 381-392 (2011).
17. Gabriel, R. et al. An unbiased genome-wide analysis of zinc-finger nuclease specificity. *Nat Biotech* **29**, 816-823 (2011).
18. Travers, K.J., Chin, C.-S., Rank, D.R., Eid, J.S. & Turner, S.W. A flexible and efficient template format for circular consensus sequencing and SNP detection. *Nucleic Acids Research* **38**, e159-e159 (2010).